DOCUMENT RESUME

ED 375 652 FL 022 492

AUTHOR Labbe, Raul; Varas, Amanda
TITLE Specialized Content Material.
PUB DATE Apr 93
NOTE 25p.; Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages (27th, Atlanta, GA, April 13-17, 1993).
PUB TYPE Reports - Descriptive (141) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS College Students; *Computer Assisted Instruction; Computer Software Development; Curriculum Design; Engineering Education; *English (Second Language); *English for Special Purposes; Foreign Countries; Higher Education; *Instructional Development; Instructional Materials; *Reading Instruction; Second Language Instruction; *Sentence Structure; Transformational Generative Grammar

ABSTRACT

A computer-based set of instructional materials, designed to teach English second-language reading skills to university-level engineering students, is described and its development is outlined. The programmed units are designed to be used individually on the computer. The stages of software development included: selection of sentence units from different registers in engineering, with attention to frequency and density of lexical items and syntax; item analysis; analysis of theoretical support in the research on syntax, semantics, written discourse, and psychology; and creation and validation of instructional units. A brief analysis is given of the unique nature of noun phrases in specialized text, which have significantly more pre- and post-modifiers clustering around nuclei than equivalent phrases in nonspecialized texts. Treatment of this complex phenomenon in the instructional design is discussed, using the approach of generative transformational grammar. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 375 652

SPECIALIZED CONTENT MATERIAL

Raúl Labbé
Amanda Varas

In the written discourse observed in specialized texts, the development of different disciplines in the realms of technical and scientific knowledge has provoked a high degree of conceptual sophistication and hence a substancial increase in syntactic complexity in ESP literature. In this area the majority of people tend to believe that the complexity of the text lies on the lexis rather than on the syntax. Obviously, for the specialized reader using a foreign language as a media, the difficulties do not arise precisely from the complexities of the vocabulary, mostly of Latin or Greek origin and of a high frequency of occurrence in their fields, but mainly from the variety of the syntactic forms which are required to convey the complex concepts that scientists usually deal with.

This research is intended to find data and criteria for the implementation of an instructional design via computer in order to develop reading skills in university students training to become engineers; reading efficiency being of paramount importance in order to have a direct access to specialized literature.

FL022492

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Raul Labbe
Amanda Varas
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

The instructional design is based on the following hyphotesis:

> "In the development of reading skills in a foreign language, the complementary application of an instructional design, via computer, is more effective than the traditional methologies so far used in the classroom".

The aforementioned design consists of an instructional package of programmed units to be used individually by the students via computer.

It is assumed that this material will present the following advantages:

1. Accessibility for the students
2. Adaptability to students needs and learning rates
3. Motivational effects characteristic of individualized learning through computers.

On the first stage the focus has been on the selection of a sample consisting of a sufficient number of sentence

units randomly selected from different registers in the area of engineering. The analysis has been done in order to determine objectively the frequency and density of lexical items and the formal devices signaling structural meaning, that is, the syntax, which would be necessary to program for the instructional design, intended to teach the students how to deal with lexis and structure when reading specialized literature.

On the second stage, the effort has been centered on the analysis and evaluation of the students' competence so as to determine those items that may interfere more drastically with the proper understanding of the reading text they are exposed to.

The third stage has been mainly devoted to investigate in the areas of syntax, semantics, written discourse and psychology which constitute the theoretical support for this research.

The fourth stage should be centered on the actual making up of the units based on the data obtained from the analysis of the specialized texts, the students' linguistic competence and the theoretical issues in psycholinguistics.

As far as the first stage is concerned, the results have clearly shown that specialized texts are relatively more complex than nonspecialized texts. This is mainly due to the fact that noun phrases in specialized texts have a significantly larger number of premodifiers and post modifiers clustering around nuclei than equivalent noun phrases in nonspecialized texts.

These findings are not surprising since they had already been accounted for by Randolph Quirk, Greenbaun, Leech and Svartvik in their well-known text "A Grammar of Contemporary English". "Scientific writing differs greatly from the other styles in having a distinctly higher proportion of noun phrases with complexity... Even a coarse-grained comparison makes clear how sensitive the noun phrase is as an index of style and how responsive it can be to the basic purpose and subject matter of any discourse". [1]

The following diagram constitutes an attempt to illustrate how the nuclei can be pre and post modified in a theoretically infinite way through the process of recursiveness. It is intended to show the different components clustering around the nucleous of the noun phrase. The diagram designed to account for pre and

postmodification items at the level of surface structure indicates the following facts:

1.- Premodifiers consist of closed system items and open system items.

2.- Closed system items are arranged in fixed distributional patterns and they seem to be irrelevant in an analysis of syntactic complexity which is generally associated to creativily expressed through recursive processes.

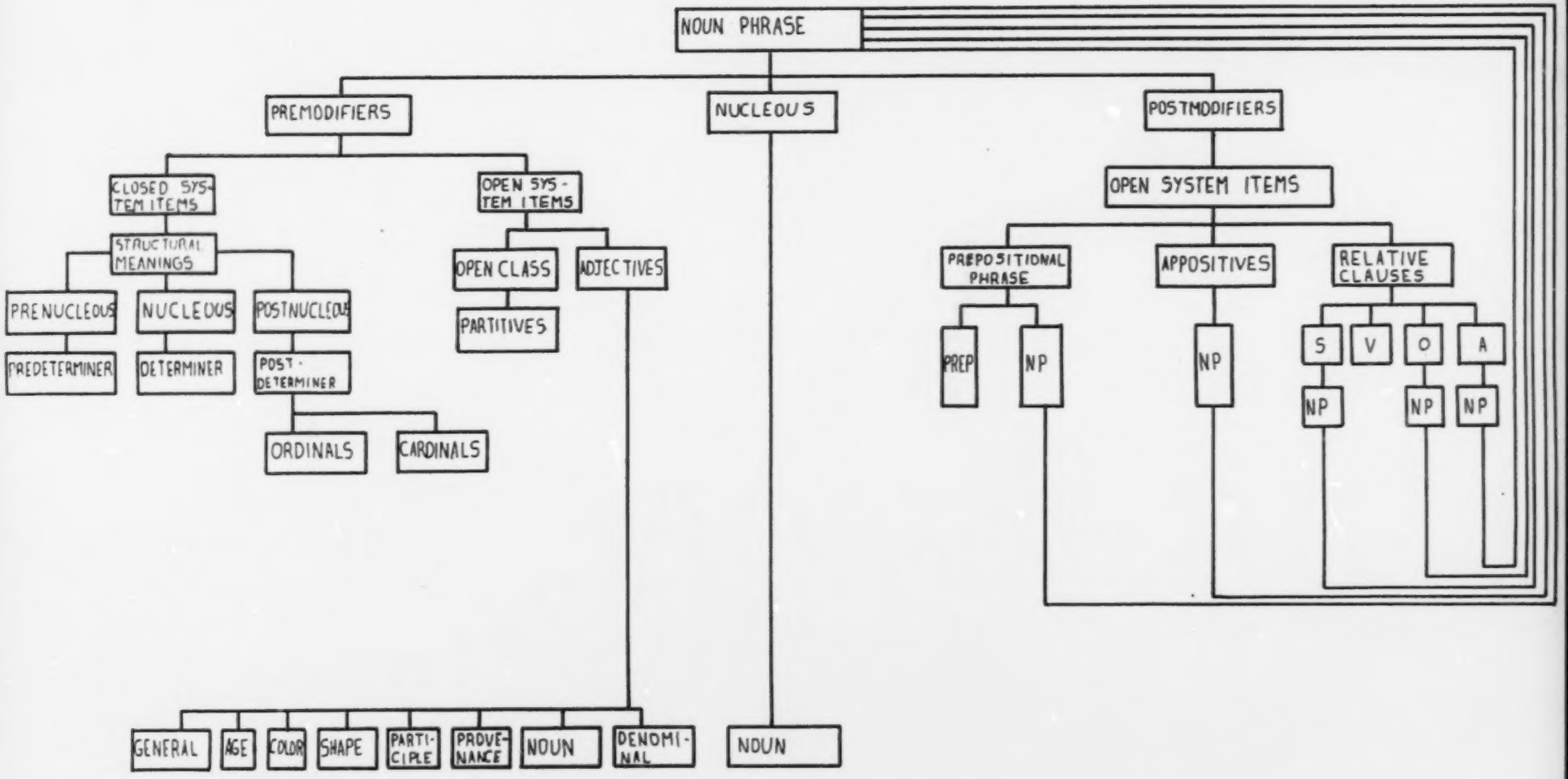
3.- Open system items mainly realized through categories like adjectives of different sorts, present participles, past participles, and nouns performing adjectival functions in the slot immediately preceding the noun acting as nucleous of the noun phrase.

4.- Postmodifers consist of prepositional phrases, appositives, and relative clauses.

5.- Noun phrases, appositives, and relative clauses may, in turn, contain all of the items which have been mentioned before through processes of recursiveness. They may be accounted for in generative transformationel terms by the formula

$$NP \rightarrow Det + Noun + N^{o} + S$$

6.- Noun phrases in prepositional phrases are also subject to recursive processes.


NOUN PHRASE
PREMODIFIERS
NUCLEOUS
POSTMODIFIERS
CLOSED SYSTEM ITEMS
OPEN SYSTEM ITEMS
OPEN SYSTEM ITEMS
STRUCTURAL MEANINGS
OPEN CLASS
ADJECTIVES
PREPOSITIONAL PHRASE
APPOSITIVES
RELATIVE CLAUSES
PRENUCLEOUS
NUCLEOUS
POSTNUCLEOUS
PARTITIVES
PREP
NP
NP
S
V
O
A
PREDETERMINER
DETERMINER
POST-DETERMINER
NP
NP
NP
ORDINALS
CARDINALS
GENERAL
AGE
COLOR
SHAPE
PARTICIPLE
PROVENANCE
NOUN
DENOMINAL
NOUN

In order to detect syntactic forms typical of the written discourse used in specialized English, it was necessary to design a notational system that would, first, show in a precise and consistent way all of those items that needed counting and, secondly, to leave out all of those items that did not need counting. The notations[8] indicate at a glance forms such as adjectives, prepositional phrases, finite and non-finite clauses, appositives, which are precisely the indexes that allow writers to incorporate new elements within sentence units.

The following notations are some of the symbols which were used in order to reduce sentences to relevant formulas, other than G.T. grammar tree diagrams, that may show just those items that need counting.

FC	_____	FINITE CLAUSE
NFC	_____	NON FINITE CLAUSE
[ ]	_____	SUBJECT, VERB, COMPLEMENT BOUNDARY
/ /	_____	SENTENCE BOUNDARY
( )	_____	PREPOSITIONAL PHRASE BOUNDARY
( )	_____	PREMODIFYNG ADJECTIVE BOUNDARY
( )	_____	SUBORDINATE CLAUSE BOUNDARY
+	_____	COORDINATING CONJUNCTION PROFORM (AND, OR, BUT)

* _____ NUCLEOUS

// _____ APPOSITIONS

Specially designed formulas, instead of G.T. tree diagrams, have been used because the interest is focused on capturing those items occurring in pre and post modifying positions rather than on describing the internal structure of "deep" constituents or the mechanics of the transformations required to arrive at the complex surface forms ordinarily used in specialized texts.

Under the general transformational approach to grammatical analysis an ordinary sentence extracted from a specialized register in the area of metal mechanics might be the following.

> 1. The foam blanket which results when ionic materials are added to a pickle bath is created by the evolution of hydrogen from the action of the acid on the steel.

For the transformational grammarian the basic goal of linguistics is to account for what the speakers know about the language. Hence a grammar is an attempt to make

explicit and conscious what the speaker of English does intuitively and unconsciously.

In connection with the sentence under analysis one of the basic tenets of G.T. grammarians is that all surface sentences consist of one or more elementary sentences. Thus, all surface sentences can be broken down into a series of elementary units, each of which consists of a noun phrase, an auxiliary, and a verb phrase.

Sentence №1 can be broken down into the following elementary sentences.

1.1 The blanket consists of a foam.

1.2 The blanket results from an action sometime some how.

1.3 Someboy adds material to a bath.

1.4 Hydrogen evolves from an action.

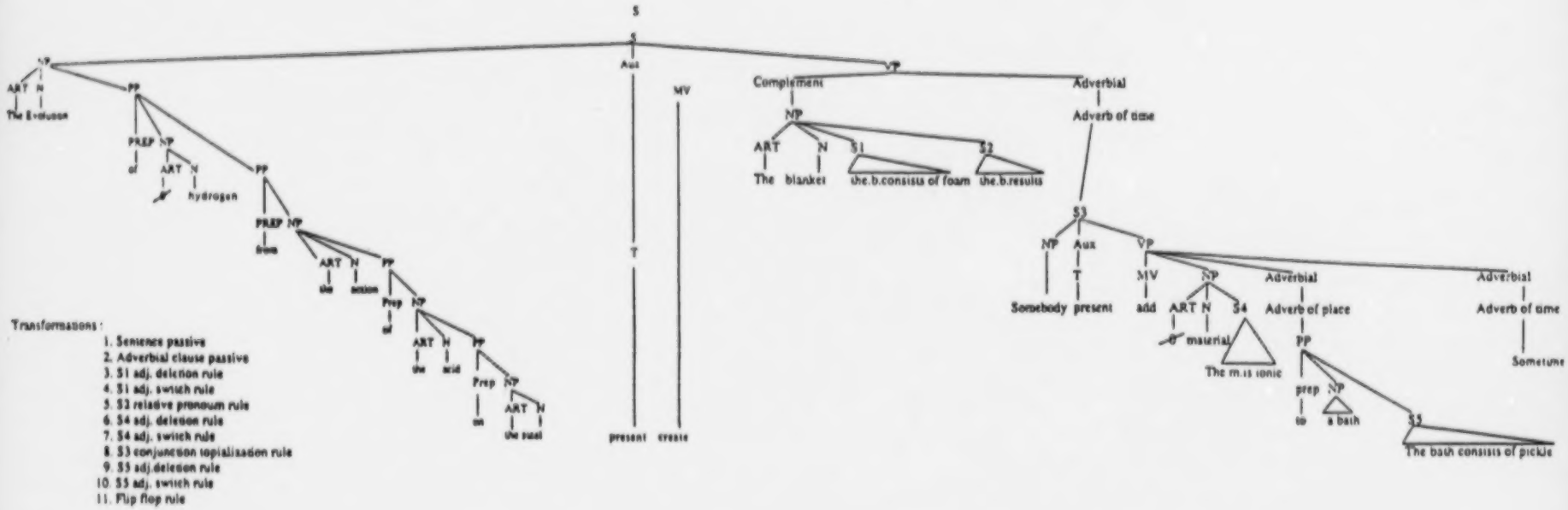
1.5 The acid acts on the steel.

1.6 An evolution creates a blanket.

1.7 The materials are ionic.

1.8 The bath consists of pickle

All of these elementary sentences should be reflected on the standard tree diagram postulated by most transformationalists, as shown below:



Transformations :

1. Sentence passive
2. Adverbial clause passive
3. S1 adj. deletion rule
4. S1 adj. switch rule
5. S2 relative pronoun rule
6. S4 adj. deletion rule
7. S4 adj. switch rule
8. S3 conjunction topicalization rule
9. S5 adj.deletion rule
10. S5 adj. switch rule
11. Flip flop rule

12

13

12

BEST COPY AVAILABLE

Since the sentences the reader deals with when exposed to a piece of specialized literature are rather complex in their structure and contain a significant number of transformations; the more complex the sentence the more transformations involved, it has been decided to adapt standard tree diagrams into formulas which may better allow the researcher to simply count those items that are responsible for syntactic variety, that is, complexity in the noun phrase. The following ad hoc formula captures all the information that is needed to account for syntactic complexity at a surface level:

$$/\ [(\text{NOUN})\ *\ \{FC_1\ \{FC_2\}\}]\ [*]\ \overset{S}{[}\ \overset{V\ \text{Passive}}{*\ \langle\ *\ \langle\ *\ \langle\ *\ \langle\ *} \overset{O}{\rangle\rangle\rangle\rangle\rangle]}\ /$$

$$FC_1\quad \overset{S}{\{[*]}\ \overset{V}{[*]}\ \}$$

$$FC_2\quad \{[(\text{adj.})*]\ \overset{S}{[*]}\ \overset{V\ \text{Passive}}{[(*)]}\ \overset{O}{\}}$$

The subject of the surface structure consists of the noun phrase "The foam blanket which results when ionic materials are added to a bath", which consists of a nucleous "blanket" preceded by a) an adjective "foam" which is really a noun adjunct, that is, a noun modifying another noun, and

b) followed by two finite clauses.

1. "Which results" and

2. "When ionic materials are added to a bath".

Finite clause 1 consists of a non-modified subject and a non-modified predicate: "which" and "results", respectively.

Finite clause 2, of a subject consisting of a nucleous "materials" premodified by the adjective "ionic", and of a non-modified predicate "are added to a bath".

The predicate of the surface structure consists of the verb phrase "is created by the evolution of hydrogen from the action of the acid on the steel".

The main verb "is created" is not regarded as relevant since there seems to be no difference in the degree of syntactic complexity shown by verb phrases in the different types of registers, in either specialized or non-specialized texts. Hence no symbols or AUX components such as tense, modal and/or aspect are included.

The agent "by the evolution of hydrogen from the action of the acid on the steel" consists of a nucleous "evolution" postmodified by four prepositional phrases "of hydrogen", "from the action", "of the acid" and "on the steel". The

degree of recursiveness is shown by the distribution of the angular brackets. Since there are four embedded prepositional phrases there are five brackets graphically showing the number of recursive rules being applied

[(* (* (* (* (*)))))]

The overall analysis allows us to count those items that contribute to syntactic complexity and to ignore those items that do not necessarily contribute to syntactic complexity.

The notational system which is postulated accounts for the surface structure of the sentence, and the notational system used by G.T. grammarians in their typical tree-branch diagrams account for both deep and surface structures. The latter analysis implies that the complexity of the sentence is the result of the number of transformations and embeddings used by the writer of the sentence to convey the ideas that he wants to put across to the reader: This being a fascinating area of research which might justify a tentative hypothesis stating that the more dense and compact the concept being expressed; the more transformations, embeddings and recursive rules being needed to generate a well formed sentence. Transforming a finite relative clause

into a nonfinite relative clause and then into a past participle are further transformations that make a particular register more compact.

From the point of view of the specific goal in this research we decided that the surface notations were more appropriate since they efficiently showed the modification structure we were interested in counting.

Regarding students competence in reading specialized English it was confirmed that the greatest difficulties did not lie on the vocabulary items but mainly on the network of formal devices that constitute the grammar of English. There seem to be no clues that may show students the mechanics of word order or distributional patterns; neither are there any clues that might suggest the reader the structural meanings of function words like "hence, although, in spite of, which, would, etc.", nor can they easily overcome hurdles like affixation or inflections posed by foreign language patterns to the reader who is trying to decode its system.

The structural items or formal devices mentioned above consist of distributional patterns such as word order;

function words such as conjunctions, auxiliaries, modals, prepositions, etc.; and inflections such as suffixes and prefixes. They allow the reader to grasp notions and meanings which are not so evident as those signaled by lexical items. These meanings refer mainly to syntactic perceptions such as actor, receiver, beneficiary, agent, tense, aspect, modality, embedding, cohesion and coherence. Without these meanings it is not possible to have access to the contents of a reading text, even if the reader may posses a perfect understanding of the lexical items contained in it.

The results of the linguistic analysis obtained on the second stage clearly indicate that the selection, grading and design of the programmed units had to be mainly based on a systematic handling of the linguistic items just referred to. In other words, the units, must be syntax centered rather than vocabulary centered.

The third stage dealing with written discourse and psychology has been focused on reading comprehension, understood as the capacity to extract the required information from a written text as efficiently as possible.

Reading involves a variety of skills including deducing the meaning and using unfamiliar lexical items, understanding explicit and unexplicit information, and mostly understanding syntactic relations within the sentence. Besides, modern studies in the area of discourse analysis have shown that a text is not a succession of separate sentences thematically related but a structure of longer units such as the paragraph or a whole text. This notion entails that the efficient reader should start with global understanding and move towards detailed understanding rather than working the other way around. This activity implies development of such skills as "understanding relations between sentences of a paragraph through lexical cohesion devices, understanding cohesion between parts of a text through grammatical cohesion devices, interpreting texts by going outside them, recognizing indicators in discourse, identifying the main point or important information in a piece of discourse, distinguishing the main idea from supporting details, skimming and scanning to locate specifically required information".[3]

The fourth stage deals with the actual making up of the programmed units designed to be used by the student on a computer. This activity has led computer scientists,

BEST COPY AVAILABLE

educators, linguists and also nonspecialists to find out how far the computer can go in interpreting natural language as opposed to artificial languages such as programming languages.

So far the potential rises of natural language processing are front-end operations. According to Noami Sager the methodology to represent the linguistic data through the computer is composed of:

1. A set of categories and sub-categories for the words of the language and a lexicon in which each word is assigned the categories and subcategories that apply to it.

2. A grammar of the language using the same categories and subcategories as the lexicon and whatever constructs are required by the type of grammar used to specify the well formed sentence structure of the language.

3. An analysis procedure or parsing program whose inputs are the text to be processed, the lexicon, and the grammar; and whose output for each sentence is a structural description of the sentence showing how the sentence was constructed according to the rules of the grammar.

So far this task has not yet been accomplished by experts in the field. Interpreting natural languages by computers still belongs to the area of science fiction.

In order to carry out our project we have had to center on a particular type of software called ADROIT (applied Data Research ON-LINE INTERACTIVE Training), which has been designed to construct a program to assist in the teaching of any subject matter (CAI), including foreign languages. In this program the student can have access to lessons previously designed for him and interact with the computer responding to the stimuli appearing on the screen.

The last stage intended to validate the programmed units will entail the following steps:

a) To expose three experimental groups of 35 students. Each to a set of computer programmed units.

b) To expose, to a traditional methodology, using the same materials prepared via CAI three control groups of 35 students per group.

BEST COPY AVAILABLE

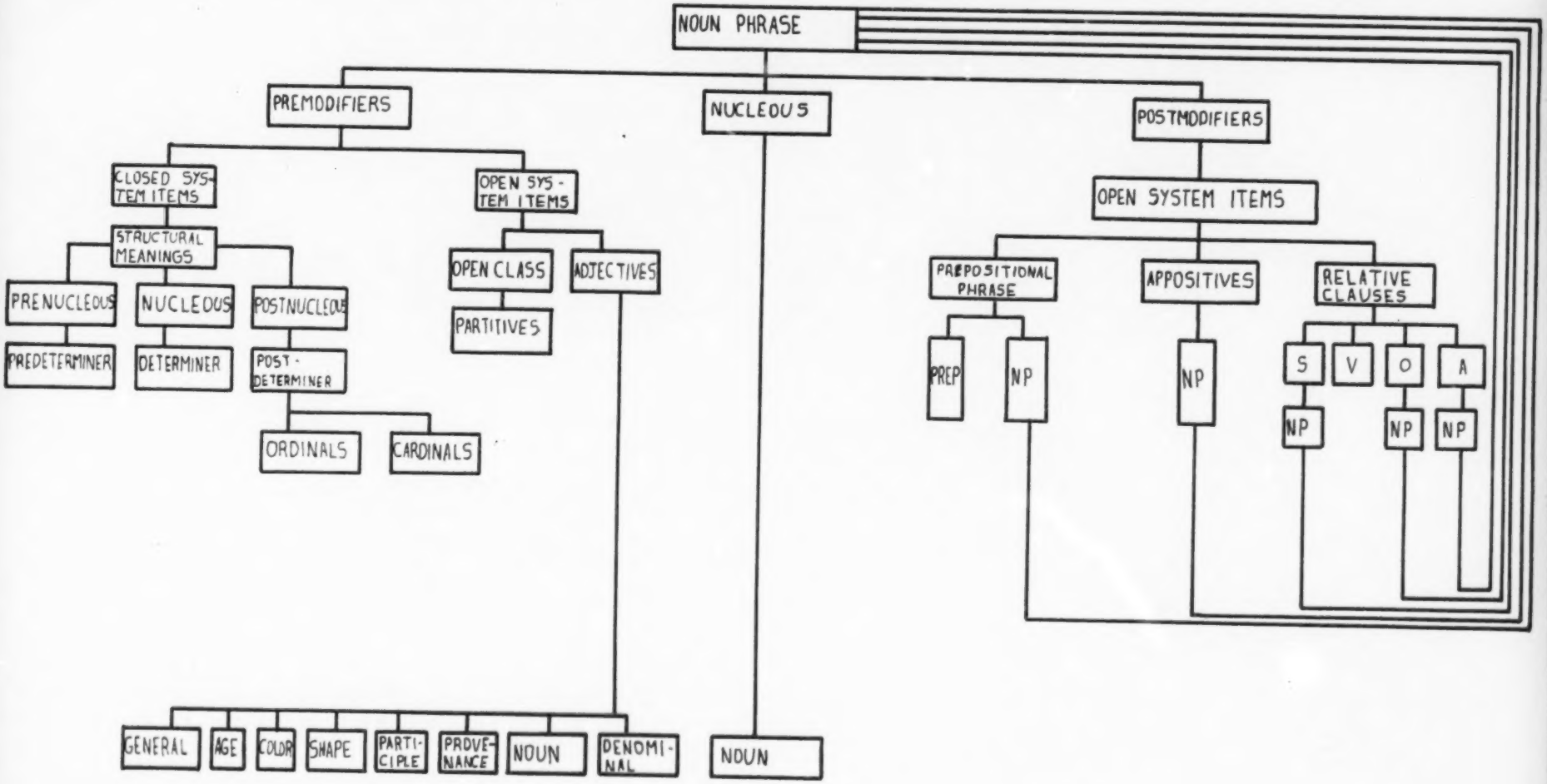
c) The results hereby obtained will be contrasted to determine through a statistical analysis if the hypothesis is supported by the facts.

BIBLIOGRAPHY

1. Randolph Quirk and others. "A grammar of Contemporary English", Longman Group Ltd., 1980, p. 934.

2. Avalos Quero Carmen "Análisis sintáctico del inglés especializado en el registro escrito". Tesis en desarrollo para optar al grado de Magister Artium en Linguistica, USACH, 1991.

3. Naomi Sager "Natural Language Information Processing". A computer grammar of English and its Applications, Addision-Wesley Published Co., 1981.


NOUN PHRASE
PREMODIFIERS
NUCLEOUS
POSTMODIFIERS
CLOSED SYSTEM ITEMS
OPEN SYSTEM ITEMS
OPEN SYSTEM ITEMS
STRUCTURAL MEANINGS
OPEN CLASS
ADJECTIVES
PREPOSITIONAL PHRASE
APPOSITIVES
RELATIVE CLAUSES
PRENUCLEOUS
NUCLEOUS
POSTNUCLEOUS
PARTITIVES
PREP
NP
NP
S
V
O
A
PREDETERMINER
DETERMINER
POST-DETERMINER
NP
NP
NP
ORDINALS
CARDINALS
GENERAL
AGE
COLOR
SHAPE
PARTICIPLE
PROVENANCE
NOUN
DENOMINAL
NOUN